

## 4. Monitoring and Evaluation of Energy Use and GHG Emissions

In Figure 4, we present an overview of the approach used in LBNL's MERVC guidelines for evaluating changes in energy use and emissions. During the monitoring and evaluation stage, gross energy savings are first measured, using one of the options provided in the U.S. Department of Energy's (DOE) International Performance Measurement and Verification Protocol (IPMVP) (Section 4.2.9). The baseline is also re-estimated, accounting for free riders (Section 4.13.1). The net change in energy use is equal to the gross change in energy use minus the re-estimated baseline. Net emissions are then calculated, using either default emission factors or emissions based on generation data (as mentioned in Section 1.4, we are only examining CO<sub>2</sub> impacts).

During the implementation of the project, monitoring of project activities is conducted periodically to ensure the project is performing as designed. We expect most, if not all, of the monitoring and evaluation activities to be performed by project developers and their contractors.<sup>1</sup> While the project is being implemented, however, we expect periodic (e.g., annual) reviews by third-party verifiers (to avoid conflicts of interest), leading to certification (see Sections 6 and 7). These verifiers might be the same independent reviewers who assessed the project proposal at the registration stage (personal communication from Johannes Heister, The World Bank, Jan. 12, 1999). As noted in Section 6, verification of energy savings and carbon emissions would be performed at certain intervals during the time the project is scheduled to save energy.

This section introduces some of the basic data collection and analysis methods used to estimate changes in energy use and associated impacts. The methods vary in cost, accuracy, simplicity and technical expertise required. Tradeoffs will need to be made for choosing the appropriate methods: e.g., level of accuracy and cost of data collection.

---

<sup>1</sup> An alternative approach is to require only certified professionals to conduct the monitoring and evaluation, as required when institutions of higher education enter into energy performance-based contracts in Texas (Texas Higher Education Coordinating Board et al. 1998). Moreover, a "professional engineer stamp" is required: (1) to certify that the monitoring and evaluation plan complies with the Texas guidelines, (2) by the person that creates the plan, (3) by the person that does the audit and cost engineering, and (4) for the person that does an independent review of the project (personal communication from Jeff Haberl, Texas A&M University, Dec. 30, 1998).

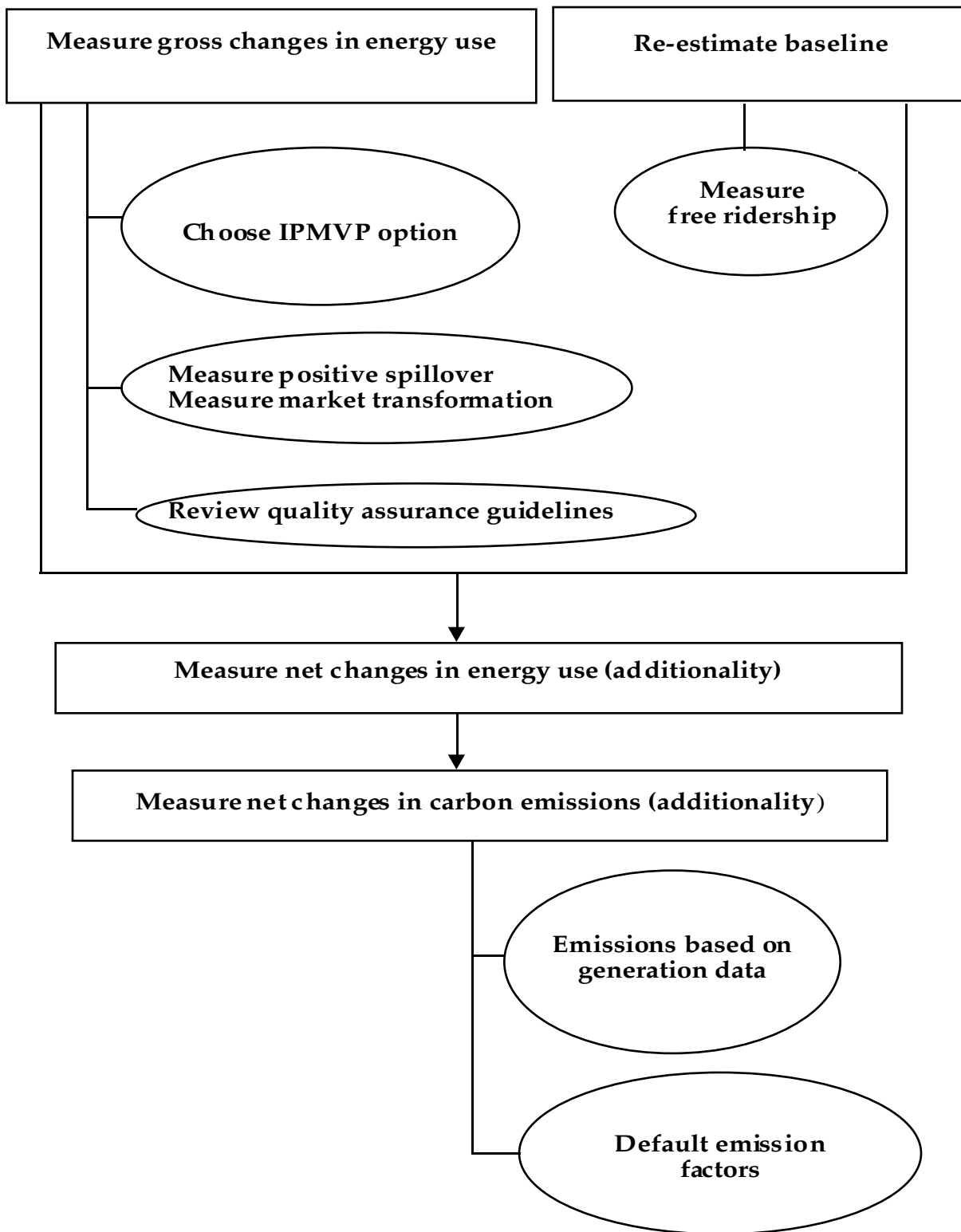


Fig. 4. Evaluation Overview

This section introduces some of the basic data collection and analysis methods used to produce energy-saving estimates (see USDOE 1994b; Raab and Violette 1994). As noted in Section 1.4, these methods have been used extensively in the evaluation of energy-efficiency programs in North America (particularly in California, the Pacific Northwest, Wisconsin, New England, and the mid-Atlantic states) (see Box 3). These methods have also been used in the evaluation of energy-efficiency programs in other countries (Hebb and Kofod 1998; Vreuls and Kofod 1997; Vine 1996a). Finally, some of the methods may be more applicable to the monitoring and evaluation of a particular project (e.g., a retrofit of a large commercial building), rather than the monitoring and evaluation of a program that involves many projects at multiple facilities (sites). If the focus is on one building, then some of the methods contained in this report will not be utilized (e.g., basic statistical models, multivariate statistical models, and some integrative methods). In the text, we indicate where these methods are appropriate for only groups of buildings; otherwise, the methods are appropriate for all situations.

Energy service companies (ESCOs) are currently using these methods in energy performance contracting. An ESCO is a company that is engaged in developing, installing and financing comprehensive, performance-based projects, typically 5-10 years in duration, centered around improving the energy efficiency or load reduction of facilities owned or operated by customers (Cudahy and Dreessen 1996; Fraser 1996). Projects are performance-based when the ESCO's compensation, and often the project's financing, are meaningfully tied to the amount of energy actually saved, and the ESCO assumes the risk in linking their compensation directly to results. Monitoring, evaluation and verification are built into the contract between the ESCO and the customer. Until recently, energy performance contracting has typically been implemented at one facility (e.g., a large commercial or industrial facility), in contrast to demand-side management projects which often promote the installation of energy-efficiency measures in many buildings (e.g., efficient lighting among residential households, chillers among hospitals, etc.). In the last few years, utilities in New Jersey and California have offered "standard performance contract" programs (pay-for-performance energy-efficiency incentive programs), resulting in energy performance contracting being conducted at multiple facilities (Goldman et al. 1998; Rubinstein et al. 1998).

**Box 3****The Evaluation of Energy-Efficiency Programs in California**

California is widely recognized as the state having the most experience in evaluating utility energy-efficiency programs in the U.S. as well as having rigorous measurement and evaluation protocols (CPUC 1998). The protocols and procedures were developed in response to the shareholder earnings mechanisms established for the four largest investor-owned utilities to acquire demand-side resources. Since 1994, the California utilities have completed hundreds of evaluation studies; earnings claims for 1994 programs and beyond have been based on adopted ex-post agreements identified in the protocols. These utilities, along with eight additional organizations, comprise the California Demand Side Management Measurement Advisory Committee which was established by the California Public Utilities Commission (CPUC) to oversee the demand-side management measurement and evaluation activities of these utilities.

The utility program evaluations have been conducted by utility staff or contractors to the utilities. The results from these evaluations are then filed with the CPUC. The CPUC's Office of Ratepayer Advocates (ORA) reviews these studies, the claimed shareholder earnings, and proposed changes or additions to the protocols. Two types of review are conducted by ORA: (1) verification of participation: a review of the utility's files to make sure all participants are in the utility's data base, and a review of the files for a random sample of participants (in some cases, onsite visits are conducted on a small sample of nonresidential customers); (2) for the larger programs, ORA prepares "review memos" that are based on a review of the evaluation studies: if problems are encountered, utility data files are requested for conducting a "replicate analysis". If ORA cannot replicate the utility analysis, then ORA will challenge the utility's results. If ORA can replicate the utility's analysis but there are problems, then more information is requested and more analyses are conducted. If ORA can replicate the utility's analysis and it is reasonable, then there is no basis for challenging the utility's results. At the end of each year, ORA files a report with the CPUC which contains recommendations on the utility evaluation studies and findings. A case management process is then conducted to see if the differences between the ORA and the utilities can be resolved. If not, then hearings are held at the CPUC to resolve the differences. At the end of the process, the Administrative Law Judge at the CPUC issues a decision on the utilities' earning claims and associated evaluation studies (where appropriate).

The California experience in measurement and evaluation is regarded by many observers to be an experience that other states (or countries) should not replicate because of the extended regulatory processes and the level of resources needed to participate in the process. However, for States (or countries) that choose to rely on utilities to promote energy efficiency as a least-cost resource with the combined set of regulations associated with Integrated Resource Planning (shareholder incentives, program cost recovery, lost revenue protection, etc.), something like the California experience is probably necessary. While the final evaluation methods and findings are clearly the best standard for the industry, nobody has made a systematic and comprehensive assessment of the costs and benefits of conducting this type of evaluation process compared to a less rigorous evaluation process. The costs are probably relatively high, but may decrease over time as the methods and their use become better known. Also, the costs are necessary to ensure that utility claims of avoided supply-side additions and shareholder incentives are reasonable.

## 4.1. Methodological Issues

Prior to reviewing the data collection and analysis methods used for measuring gross and net energy savings and GHG emissions, we first discuss two key methodological issues: measurement uncertainty, and the frequency and duration of monitoring and evaluation. These issues are not only addressed in the monitoring and evaluation stage but should also be examined in the project design stage.

### 4.1.1. Measurement uncertainty

While there are several types of uncertainty that can affect the actual realization of GHG reductions, uncertainty in the measurement of GHG reductions needs to be taken into account when presenting monitoring and evaluation findings.<sup>1</sup> Measurement uncertainties include the following: (1) the use of simplified representations with averaged values (especially emission factors); (2) the uncertainty in the scientific understanding of the basic processes leading to emissions and removals for non-CO<sub>2</sub> GHG; and (3) the uncertainty in measuring items that cannot be directly measured (e.g., project baselines). Some of these uncertainties vary widely by type of project (depending on approach, level of detail, use of default data or project specific data, etc.), and length of project (e.g., short-term versus long-term). It is important to provide as thorough an understanding as possible of the uncertainties involved when monitoring and evaluating the impacts of energy-efficiency projects.

---

<sup>1</sup> Other types of uncertainty include the following: (1) project development and construction uncertainty, i.e., the project won't be implemented on time or at all, even though funds have been spent on project development; (2) operations and performance uncertainty (e.g., if the energy-consuming equipment is not used as projected, then carbon savings will change); and (3) environmental uncertainty (IPCC 1995; USAID 1996; UNFCCC 1998b). Project developers should provide a description of the project developer's experience, existing warranties, the reputation of equipment manufacturers, the performance history of previous projects, and engineering due diligence. The political and social conditions that exist that could potentially affect the credibility of the implementing organizations (e.g., political context, stability of parties involved and their interests, and potential barriers) also need to be described.

Because of the difficulties and uncertainties in estimating energy savings and reduced emissions, the level of precision and confidence levels associated with the measurement of savings need to be identified.<sup>1</sup> Project developers and evaluators should report the precision of their measurements and results in one of two ways: (1) quantitatively, by specifying the standard deviation around the mean for a bell-shaped distribution, or providing confidence intervals around mean estimates; or (2) qualitatively, by indicating the general level of precision of the measurement (e.g., low, medium or high).

It is unclear at this time on how uncertainty will be treated in the calculation and crediting of energy savings and reduced emissions. At a minimum, the most conservative figures should be used at every stage of calculation (e.g., the lower boundary of a confidence interval). The qualitative assessment of uncertainty is more problematic, however, some type of discounting or debiting could be used to adjust energy savings and reduced emissions in situations where there is a great deal of uncertainty. Where there is substantial uncertainty, project developers need to design higher quality energy-efficiency projects so that impacts are more certain.

In conclusion, the evaluation of energy-efficiency projects should: (1) evaluate the project's contingency plan, where available, that identifies potential project uncertainties and discusses the measures provided within the project to manage the uncertainties; (2) identify and discuss key uncertainties affecting all emission estimates; (3) assess the possibility of local or regional political and economic instability and how this may affect project performance; and (4) provide confidence intervals around mean estimates.

#### **4.1.2. Frequency and duration of monitoring and evaluation**

The frequency of monitoring and evaluation will most likely be linked to the schedule of transfer of carbon credits.<sup>2</sup> It is possible that these credits could be issued on an annual basis. The frequency of monitoring and evaluation will also depend on the variables being examined and methods used: e.g., hourly end-use monitoring conducted for a two-week period, or short-term monitoring of lighting energy use for five-minute periods. The monitoring period may last longer than the project

---

<sup>1</sup> Unless otherwise noted, we assume normal distributions, represented by a normal, bell-shaped curve in which the mean, median and mode all coincide.

<sup>2</sup> Other models are possible (e.g., up-front lump-sum payment), but unlikely since the issuance of certified emission reduction units occurs after a verification process.

implementation period: for example, a project to install compact fluorescent lamps may last 3 years, but electricity savings from those lamps will continue beyond the three years.

The **persistence** of the energy savings from energy-efficiency projects is a critical issue in the monitoring and evaluation of energy savings, as well as in the design and implementation of the projects. The institutional, community, technical and contractual conditions likely to encourage persistence are of utmost concern. In some cases, encouraging the participation of community members in the development and implementation of energy-efficiency projects will help to ensure the longevity of a project, although the design and implementation process may take longer and costs will increase. Project persistence will also increase by encouraging operations and maintenance, providing spare parts and equipment, and making sure technical expertise is available. Finally, contracts can incorporate provisions that lead to debiting of emission reduction units (for the host and/or investor country) if a project does not last as long as expected.

The issue of persistence is directly linked to the concept of market transformation (Section 3.1.3). Markets are transformed as market barriers are reduced due to market intervention. The reduction in market barriers is reflected in a set of market effects that last after the market intervention has been withdrawn, reduced or changed. For example, an energy-efficiency project may reduce awareness barriers by providing information to a targeted audience (e.g., building owners and managers). The key question for market transformation (and persistence) is whether the targeted audience remains informed once the project has ended: if there is no persistence, then there is no market transformation; if there is some persistence, then market transformation is possible.

As mentioned at the beginning of this section, energy service companies conduct energy performance contracting in one or more buildings, and their compensation, and often the project's financing, are tied to the amount of energy actually saved. Because the persistence of energy savings is of paramount interest for all concerned, periodic (if not continuous) monitoring and evaluation is built into the contract between the ESCO and the customer. For example, when institutions of higher education in Texas enter into energy performance-based contracts, they require periodic monitoring to guarantee the energy savings in their contract (Texas Higher Education Coordinating Board et al. 1998).

In California, investor-owned utilities must periodically conduct two types of persistence studies on energy-efficiency measures: retention studies and performance studies (CPUC 1998). The retention studies assess the fraction of measures installed in the first program year which are still in place and operable at the time of the study. The data are collected by telephone, on-site or mail surveys from program participants. In the performance studies, the performance/efficiency of the equipment is measured on site; the studies are conducted every four or five years.

The U.S. Environmental Protection Agency's (EPA) Conservation Verification Protocols (CVP) contains disincentives to encourage monitoring over the life of the measure (see Section 1.6.3). Three options are available for evaluating subsequent-year energy savings (Table 2): monitoring, inspection and a default (Meier and Solomon 1995; USEPA 1995 and 1996). The estimated impacts of the energy-efficiency measures eligible for emissions credits are those that can be demonstrated with at least a 75% level of confidence. This means that there must be a 75% likelihood that the true level of impacts is equal to or greater than the value calculated in the evaluation (i.e., there can be no more than a 25% likelihood that actual impacts are less than those reported by the evaluation). The evaluation must be designed to produce this level of confidence in the final evaluation estimates.

**Table 2. Options for Obtaining Credit for Energy Savings Over Time**

<p><b>Monitoring option</b></p> <p>By monitoring over the life of the measure, one obtains credit for a greater fraction of the savings and for a longer period of time. Biennial verification in subsequent years 1 and 3 (including inspection) is required, and savings for the remainder of physical lifetimes are the average of the last two measurements. The monitoring option requires a 75% confidence in subsequent-year savings.</p>
<p><b>Default option</b></p> <p>By relying on default (stipulated) savings, allowable savings are restricted: credit is only for 50% of first-year savings, and limited to one-half of the measure's physical lifetime.</p>
<p><b>Inspection option</b></p> <p>By inspecting (confirming) that measures are both present and operating, credit is allowed for 75% of first-year savings and is limited to one-half of the measure's physical lifetime (with biennial inspections), or 90% of first-year savings for physical lifetimes of measures that do not require active operation or maintenance (e.g., building shell insulation, pipe insulation and window improvements).</p>

Source: Derived from USEPA (1995 and 1996)

Finally, where more than one project is being implemented, evaluators should evaluate a project by its persistence or lack of persistence — this will be reflected in “project lifetime,” which may be different than an expected lifetime of a project as initially proposed by developers. For example, if a project area is likely to undergo serious changes in 10 years, then the carbon emission reductions for

that project are limited to that 10-year lifetime. The value of those reduced emissions may be less than for emissions from similar projects that are expected to last longer (e.g., 20 years). Accompanying the evaluation, the evaluator should provide a list of indices that demonstrate the potential for persistence: e.g., type and number of income groups targeted by project, potential socioeconomic impacts addressed (see Section 8.2), local manufacturing capability, potential sources of uncertainty and risk addressed (see Section 4.1.1), etc.

## 4.2. Measurement of Gross Energy Savings

As described at the beginning of this section, the first step in measuring emission reductions is the measurement of gross energy savings<sup>1</sup>: comparing the observed energy use of project participants with pre-project energy consumption.<sup>2</sup> Several data collection and analysis methods are available which vary in cost, precision, and uncertainty. The data collection methods include engineering calculations, surveys, modeling, end-use metering, on-site audits and inspections, and collection of utility bill data. Most monitoring and evaluation activities focus on the collection of measured data; if measured data are not collected, then one may rely on engineering calculations and “stipulated” (or default) savings (as described in EPA’s Conservation Verification Protocols and in DOE’s International Performance Measurement and Verification Protocol (Section 4.2.9)).<sup>3</sup> Data analysis methods include engineering methods, basic statistical models, multivariate statistical models (including multiple regression models and conditional demand models), and integrative methods. As mentioned at the beginning of this section, the use of these methods will vary by how many buildings are being evaluated.

---

<sup>1</sup> LBNL’s MERVC guidelines focus on energy use (e.g., kWh and fuel use), and not demand (e.g., kW) because CO<sub>2</sub> emissions depend on the amount of kWh that must be supplied, not the power capacity saved.

<sup>2</sup> Takeback (or snapback or rebound) is a price effect where program participants increase their demand for energy services when efficiency measures decrease the price of services. We do not discuss takeback in this report because most researchers believe that takeback of energy savings is minimal, with the possible exception of low-income programs that affect customers who are consuming energy services below their comfort level (Violette et al. 1998).

<sup>3</sup> Stipulated savings refer to two different types of stipulated savings methods: (1) algorithms for calculating energy savings for specific measures; and (2) a set of criteria for using best-engineering practices (USEPA 1995). The rationale for the use of stipulated savings is that the performance of some energy-efficiency measures is well understood and may not be cost effective to monitor; stipulated savings should only be used for certain retrofits and conditions.

In this section, we provide a brief review of methods to provide guidance to evaluators. For each method, we provide examples of applications of these methods; the examples are for illustrative purposes. The methods used for data collection and the evaluation of non-electric end-use efficiency projects are similar to those used for electric end-use efficiency projects; there will, however, often be greater reliance on engineering methods and surveys because centralized billing information will generally not exist.

#### **4.2.1. Establishing the monitoring domain**

During the project design stage, the project developer needs to determine who will be monitored: just program participants, or nonparticipants, too. In the beginning stages of a project, the indirect impacts of a project are likely to be modest as the project gets underway, so that the MERVC of such impacts may not be a priority. These effects are also likely to be insignificant or small for small projects. Under these circumstances, it may be justified to disregard these impacts and simply focus on energy savings from the project. This would help reduce MERVC costs. As the projects become larger or are more targeted to market transformation, these impacts should be evaluated.

Currently, there are weak linkages in assessing multiple monitoring domains (e.g., local, regional and national) (Andrasko 1997). One potential solution to strengthening these linkages is the use of “nested monitoring systems” where an individual project’s monitoring domain is defined to capture the most significant energy savings and where provisions are made for monitoring energy use and carbon emissions outside of the project area by regional or national monitoring systems (Andrasko 1997).

#### **4.2.2. Engineering methods**

Engineering methods are used to develop estimates of energy savings based on technical information from manufacturers on equipment in conjunction with assumed operating characteristics of the equipment. The two basic approaches to developing engineering estimates are engineering algorithms and engineering simulation methods (Violette et al. 1991).

Engineering algorithms are typically straightforward equations showing how energy (or peak) is expected to change due to the installation of an energy efficiency measure. They are generally quick and easy to apply but are limited to certain types of retrofits (e.g., motor replacement on constant use motor). The accuracy of the engineering estimate, however, depends upon the accuracy of the

inputs, and the quality of data that enters an engineering algorithm can vary dramatically. Hence, calibration to measured data is often necessary for using algorithms.

Engineering building simulations are computer programs that model the performance of energy-using systems in residential and commercial buildings.<sup>1</sup> These models use information on building occupancy patterns, building shell and building orientation (e.g., window area, building shape and shading) and information on all of the energy-using equipment. The input data requirements for the more complex simulation models are extensive and require detailed onsite data collection as well as building blueprints (e.g., see Box 4).

Building simulation models are best suited for space heating/cooling analyses and for predicting interactive effects of multiple measure packages where one of the measures influences space conditioning.<sup>2</sup> Measures best addressed by simulation models include heating, ventilation, and air-conditioning (HVAC) measures, building shell measures, HVAC interactions with other measures, and daylighting measures. Equipment measures such as lighting, office equipment, and appliance use are typically calibrated outside the simulation, except for their interactive impacts.

Building simulation models are tools, and their usefulness is a function of the skill of the modeler, the accuracy of the input information, and the level of detail in the simulation algorithms. A key component of building energy simulation methods is the appropriate calibration of these models to actual consumption data. The calibration could involve monthly energy consumption data from bills (at a minimum), kW demand meters, run-time meters, and short-term end-use metering (e.g., two to six weeks of metering). One advantage of simulation models is that they take into account such factors as weather data and interactions between the HVAC system and other end uses. A primary disadvantage of building simulation tools is that they are very time consuming and usually require specialized technical expertise, making them costly in the long run. In addition, because they simplify processes, they may work well on average but may not necessarily work well for a particular building (or vice versa). Finally, the behavior-driven inputs (e.g., hours of operation) are often subject to self-report bias.

---

<sup>1</sup> Building energy simulations have been carried out in many countries outside of North America including: Australia (Yune 1998), Brazil (Lamberts et al. 1998), China, Hong Kong, Mexico, New Zealand, Pakistan, Saudi Arabia, Singapore, South Africa, South Korea, Sweden, and Switzerland (personal communications from Joe Huang and Fred Winkelmann, Lawrence Berkeley National Laboratory, Nov. 12, 1998).

<sup>2</sup> The simulation results can be produced in kWh, therms, or Btus. Given the fuel efficiency of the heating system, the amount of fuel required to meet the heating demand of the building can be calculated.

**Box 4****Engineering Building Simulation Example**

The Pacific Gas and Electric Company and Southern California Edison contracted with a consulting firm to perform a comprehensive evaluation of their 1994 nonresidential new construction programs. These programs offered incentives for building envelope, lighting, HVAC and refrigeration measures, with the aim of encouraging the construction of buildings more energy efficient than mandated by statewide building codes.

Evaluation methods: The gross impact analysis was conducted using the DOE-2 building energy simulation program. DOE-2 is a very flexible modeling tool that allows the calculation of energy and demand savings for lighting, lighting controls, shell measures, HVAC efficiency improvements, many HVAC control measures, and grocery store refrigeration systems. An automated process that integrated on-site data collection and DOE-2 modeling conducted DOE-2 simulations of 347 sites under multiple baseline scenarios. A DOE-2 model was constructed for each surveyed building, and the engineering analysis used Typical Meteorological Year weather data representative of the building's location.

Model calibration to billing data was used to provide a check on the model results. Calibration procedures focused on high influence parameters, such as outside air fraction, economizer operation, fan schedules, etc. that may be difficult to observe during an on-site survey. Models were calibrated to  $\pm 10\%$  agreement on monthly whole-building energy consumption, where possible.

A second round of calibrations was performed on a sub-sample of 30 sites where short-term monitored data were collected. The short-term monitoring was used to improve the end-use consumption estimates in all building models, thus improving estimates of energy savings for the entire sample. Data gathered from short-term monitoring was used to define key simulation model inputs, thus limiting the key variables available for adjustment during calibration. This ensured that building systems were modeled as they actually operated.

Evaluation concerns: (1) in collecting extensive billing data, the study was delayed and may have done more harm than good: only a fraction of the billing data proved to be useful and had a relatively small impact on the results, while the delay made surveying decision makers and obtaining permission for on-site audits more difficult; (2) the use of a commercial database as a sample frame led to ambiguities in the identity and location of program participants; and (3) the collection of building standard documentation was frustrating as many companies viewed this documentation as proprietary and refused to release it: as a result, very little documentation was collected.

Findings: The PG&E program resulted in a gross summer on-peak demand savings of 19.7 MW and an annual energy savings of 81,350 MWH. The SCE program resulted in a gross summer on-peak demand savings of 10.3 MW and an annual energy savings of 67,850 MWH.

Source: Pacific Gas and Electric. 1997. *Impact Evaluation of Pacific Gas and Electric Company and Southern California Edison 1994 Nonresidential New Construction Programs*. March 1. San Francisco, CA: Pacific Gas and Electric.

Engineering estimates (in algorithms and building simulations) are often developed as part of an ongoing project tracking database. Because of changes during project implementation, the engineering assumptions used at the design stage of a project need to be changed as evaluation data are collected (e.g., number of operating hours and specific measures installed). Engineering methods for use in assessing the impacts of energy-efficiency projects are improving as experience points out their strengths and weaknesses. Their value for impact evaluation also is increasing as actual field data is used to adjust or recalculate savings estimates. Engineering methods are often used as a complement to other evaluation methods rather than serving as stand-alone estimates of project impacts (see below).

Although engineering approaches are improving and increasing in sophistication, engineering estimates generally produce estimates of baseline energy use and project impacts that do not account for free riders (Section 3.2.1) and positive project spillover (Section 3.1.2). It is possible to incorporate free rider and spillover factors from surveys and other evaluation sources in order to calculate more accurately baseline energy use and project impacts. Engineering analyses may be most appropriate for: (1) the initial year of project implementation where monitoring will rely on engineering estimates and where data have not been collected; (2) projects where small savings are expected (making less expensive methods preferable); (3) large industrial customers (making it difficult to find a representative comparison group of customers); (4) new construction projects (where pre-project energy use does not exist); and (5) certain types of retrofits (e.g., motor replacement for a constant use motor).

In sum, the advantages of engineering methods are that engineering algorithms are relatively quick and inexpensive to use (in contrast to building simulations that are typically more resource intensive) and are probably most useful when integrated with other data collection and analysis methods. The primary disadvantage is that the data used in the calculations rely on assumptions that may vary in their level of accuracy. Accordingly, engineering analyses need to be “calibrated” with onsite data (e.g., operating hours and occupancy). Thus, as project information is collected, engineering estimates can be improved.

**Table 3. References to Engineering Methods**

Examples	References
Residential new construction Commercial heating, ventilation & air-conditioning Commercial lighting Commercial new construction Commercial new construction Commercial retrofit Commercial retrofit Commercial retrofit Commercial energy management systems Commercial chillers Industrial process, refrigeration, and miscellaneous measures Industrial heating, ventilation & air-conditioning	Mahone et al. (1996) Baker et al. (1996) Caulfield and Galawish (1996) Sebold and Wang (1996) Carlson et al. (1997b) Katipamula and Claridge (1993) Lui and Claridge (1998) Haberl and Claridge (1985) Wortman et al. (1996) Carlson et al. (1997a) Clarke et al. (1996)  Mowris et al. (1996)
	General References
	Claridge (1998) Jacobs et al. (1992) Knebel (1983) Ridge et al. (1997) USDOE (1997) Violette et al. (1991)

#### 4.2.3. Basic statistical models for evaluation (for groups of buildings)

Statistical models that compare energy consumption before and after the installation of energy efficiency measures have been used as an evaluation method for many years (Violette et al. 1991). The most basic statistical models simply look at monthly billing data before and after measure installation using weather normalized consumption data (this is particularly important where weather-dependent measures are involved—e.g., heating and cooling equipment, refrigerators, etc.). If the energy savings are expected to be a reasonably large fraction of the customer's bill (e.g., 10% or more), then this change should be observable in the project's bills. Smaller changes (e.g., 4%) might also be observed in billing data, but more sophisticated billing analysis procedures are often required. This method can be used for comparing changes in energy use for project participants and a comparison group (e.g., see Box 5). Statistical models are most useful where many projects (or one project with many participants) are being implemented (e.g., in the residential sector).

These simple statistical comparison estimates rely on the assumption that the comparison group is, in fact, a good proxy for what project participants would have done in the absence of the project. However, there are reasons to expect systematic differences between project participants and a comparison group (e.g., participants may already be more inclined to adopt a measure than

nonparticipants do). Consequently, evaluators may start with a basic statistical approach because it is relatively inexpensive and easy to explain, but they should consider augmenting this method with survey data and other measurements to test the underlying assumptions of the model. Additional modeling and verification methods may be needed before the results of these basic comparisons can be accepted as accurately representing the actual impacts of an energy-efficiency project.

### Box 5

#### Basic Statistical Model Example (for groups of buildings)

The Ohio Department of Development's Office of Energy Efficiency contracted with a consulting firm to perform a comprehensive evaluation of the Ohio Low-income Home Weatherization Assistance Program. The program evaluation compared the energy use of participants and a comparison group, using a software model called the Princeton Scorekeeping Method, or PRISM (see Box 4). Approximately 95% of the utility participants were served by one of eight local utilities owned by 6 utility companies. A key task in the study was to collect and clean the needed data for assessing energy usage.

Evaluation methods: The data collection process began in early 1996 with the gathering of statewide weatherization databases for program years 1994 and 1995. The participant utility account numbers, recorded by local weatherization agencies, were checked and cross-referenced to other databases to create the most accurate and complete participant account lists. Energy usage was formally requested from utilities in June of 1996. The data requested included approximately 3 years of usage data.

PRISM was used to analyze the gas usage data for the 1994 low-income weatherization assistance program participants and a comparison group drawn from the 1995 participants. PRISM provides weather-adjusted annual energy consumption estimates based on monthly usage data. Savings for each house were calculated as the difference in the normalized annual consumption rates between the pre- and post-treatment periods. For the comparison group, the pre-period was defined as the period two years prior to actual treatment, and the post-period was the year immediately preceding actual treatment.

Evaluation concerns: (1) cleaning the utility usage and payment data was a major task; (2) sample attrition (usage data were acquired for just 70% of participants); and (3) usage anomalies and/or incomplete data, which led to the exclusion of 23% of the PRISM savings estimates due to unreliable or physically impossible PRISM results in either the pre or post periods.

Findings: Preliminary results indicated that the program produced impressive gas savings of more than 300 ccf/year, and 400 ccf/year for high-use households. The savings enabled low-income customers to better afford their utility service, avoiding collection actions and service disconnections.

Sources: (1) Blasnik, M. 1997. "A Comprehensive Evaluation of Ohio's Low-Income HWAP: Big Benefits for Clients and Ratepayers," in the *Proceedings of the 1997 International Energy Program Evaluation Conference*. pp. 301-308. Chicago, IL: National Energy Program Evaluation Conference. (2) Fels, M. 1986. "PRISM: An Introduction," *Energy and Buildings* 9(1-2): 5-18.

The advantages of basic statistical models are that comparing the billing data is inexpensive, and the results are easy to understand and communicate. The disadvantages include limited applicability (because of the need for stable building operations or lack of prior billing records (e.g., new construction)), participant samples of significant size are required for validity, and peak impacts cannot be evaluated.

**Table 4. References to Basic Statistical Models  
(for groups of buildings)**

Examples	References
Residential weatherization Low-income weatherization Commercial heating, ventilation & air-conditioning	Bohac et al. (1996) Blasnik (1997) Baker et al. (1996)
	<b>General References</b>
	Fels (1986) Ridge et al. (1997) Violette et al. (1991)

#### 4.2.4. Multivariate statistical models for evaluation (for groups of buildings)

In project evaluation, more detailed statistical models may need to be developed to better isolate the impacts of an energy-efficiency project from other factors that also influence energy use. Typically, these more detailed approaches use multivariate regression analysis as a basic tool (Box 6) (Violette et al. 1991). Regression methods are simply another way of comparing kWh or kW usage across dwelling units or facilities and comparison groups, holding other factors constant. Regression methods can help correct for problems in data collection and sampling. If the sampling procedure over- or under-represents specific types of projects (e.g., large-scale energy intensive projects) among either project participants or the comparison group, the regression equations can capture these differences through explanatory variables. Two commonly applied regression methods are conditional demand analysis (CDA) and statistically adjusted engineering models (Violette et al. 1991).

Some define CDA strictly as a very specific and complex regression-based approach that should include, among other independent variables, a complete inventory of all major energy-using equipment (see Ridge et al. 1994). Others define CDA less restrictively as a collection of regression-based approaches that specify energy consumption as conditional on any number of measured variables, but not a complete inventory of equipment. Statistically adjusted engineering models

would fall into this category. Most impact evaluations of energy-efficiency programs fall into the general category of non-classic, less restrictive CDA. Because of its greater data requirements, the classic, restrictive CDA model experiences greater measurement error, sample error and non-response error than a model that has less demanding data requirements. However, these same data requirements also mean that it will be less likely to omit a relevant variable. Similarly, CDA models that have much less demanding data requirements than the more restrictive CDA models will experience less measurement error, sample error and non-response error. However, these same data requirements mean that there is a greater likelihood that a relevant variable will be omitted.

### Box 6

#### **Multivariate Statistical Model Example (for groups of buildings)**

In 1982, Southern California Edison contracted with a consulting firm to conduct an impact evaluation of its commercial and industrial conservation program in which commercial and industrial customers received cash rebates for installing energy-saving devices.

Evaluation methods: In addition to an engineering analysis of energy savings, a multivariate statistical analysis of energy savings was conducted to account for variations in weather patterns and customer characteristics that affect energy consumption and realized savings. The savings estimates were based on a statistical analysis of customers' bills for a period spanning at least 1 year before and 1 year after the equipment was installed. A separate analysis was performed for each type of equipment, using only the bills for customers who installed that type of equipment. The equations included variables that were used to account for three components of consumption: baseload, weather-sensitive consumption, and the conservation effect. The variables explaining base, non-weather-sensitive load were hours of operation per month, square footage, average price of electricity, time trend indicators for the demand group of each customer, and an indicator for whether the customer was commercial or industrial. Weather sensitivity was captured by a cooling degree days variable. The effect of installing equipment under the program was captured with a dummy variable indicating that the customer had installed the equipment.

Findings: The amount of variability (adjusted R-squared) explained by these models varied by type of equipment: 0.21 for time clocks, 0.75 for photocells, 0.75 for load controllers, 0.60 for HVAC economy cycle, 0.56 for lighting system changes, and 0.74 for low wattage fluorescent lamps.

Source: Train, K., P. Ignelzi, and M. Kumm. 1985. "Evaluation of a Conservation Program for Commercial and Industrial Customers," *Energy* 10(10):1079-1088.

**Table 5. References to Multivariate Statistical Models  
(for groups of buildings)**

Examples	References
Residential new construction Residential new construction Commercial retrofit Commercial retrofit Commercial and industrial retrofit Commercial new construction Commercial heating, ventilation & air-conditioning	Mahone et al. (1996) Gunel et al. (1995) Katipamula et al. (1994) Coito and Barnes (1996) Fagan et al. (1995) Heitfield et al. (1996) Randazzo et al. (1996)
	<b>General References</b>
	Claridge (1998) Reddy et al. (1998) Ridge et al. (1997) Violette et al. (1991)

#### 4.2.5. End-use metering

Energy savings can be measured for specific equipment for specific end uses through end-use metering (Box 7) (Violette et al. 1991). This type of metering is conducted before and after a retrofit to characterize the performance of the equipment under a variety of load conditions. The data are often standardized (normalized) for variations in operations, weather, etc. The advantage of end-use metering is that it provides a greater degree of accuracy than engineering estimates or short-term monitoring for measuring energy use (Box 7) (see Section 3.3.5). End-use meters calculate the energy change on an individual piece of equipment in isolation from the other end-use loads (as opposed to billing analysis, which captures the effect at the whole building level). Hence, end-use metering reduces measurement error (assuming the metering equipment is reliable) and reduces the number of control variables required in models.

The disadvantages of end-use metering are: (1) it requires specialized equipment and expertise, typically more costly than the other methods, and therefore most samples need to be small; (2) the small samples may lead to biases in sample selection and problems in representativeness; (3) end-use metering of post-participation energy consumption alone does not, in and of itself, improve estimates of project impacts; (4) end-use metering experiments to measure both pre-and post-installation consumption are difficult to construct, especially in identifying project participants before their becoming participants to allow the pre-measure end-use metering; and (5) it cannot by itself be used to estimate free riders and positive project spillover. Accordingly, end-use metering is more often

seen as a data collection method (rather than a data analysis method) that can provide useful information for integrative methods (see Section 4.2.7).

### Box 7

#### End-use Metering Example

The Central Maine Power Company contracted with a consulting firm to conduct an impact evaluation of its residential new construction program. The program was designed to improve the energy efficiency of new homes being built in the area.

Evaluation methods: Space heating electricity use was metered. As part of the evaluation, the consultants constructed a conditional demand model using billing data for space heating only as the dependent variable. The regression model only controlled for variables that influenced space heating: e.g., use of wood heating, square footage, thermostat setback usage, presence of heated basement, R-value of ceiling and wall insulation, etc.

Findings: The amount of variability (adjusted R-squared) explained by this model was 0.72, a large R-squared given the small sample size (22 observations).

Source: Central Maine Power Company. 1990. *Evaluation of the Energy Savings Resulting from Central Maine Power Company's Good Cents Home Program*. Augusta, ME: Central Maine Power Company.

**Table 6. References to End-use Metering**

Examples	References
Residential new construction Commercial chillers Commercial chillers and motors Commercial lighting Thermal energy storage Commercial heating, ventilation & air-conditioning	Central Maine Power (1990) Carlson et al. (1997a) Quackenbush et al. (1997) Amalfi et al. (1996) Michelman et al. (1995) Dohrmann et al. (1995)
	<b>General Reference</b>
	Violette et al. (1991)

#### 4.2.6. Short-term monitoring

Short-term monitoring refers to data collection conducted to measure specific physical or energy consumption characteristics either instantaneously or over a short time period. This type of monitoring is conducted to support evaluation activities such as engineering studies, building

simulation and statistical analyses (Violette et al. 1991). Examples of the type of monitoring that can take place are spot watt measurements of efficiency measures, run-time measurements of lights or motors, temperature measurements, or demand monitoring (e.g., see Box 8). Short-term monitoring is gaining increasing attention as evaluators realize that for certain energy-efficiency measures with relatively stable and predictable operating characteristics (e.g., commercial lighting and some motor applications), short-term measurements will produce gains in accuracy nearly equivalent to that of long-term metering at a fraction of the cost.

### Box 8

#### Short-term Monitoring Example

In this example, short-term monitoring of lighting systems was undertaken in the context of an EPRI Tailored Collaboration aimed at developing short-term monitoring techniques for evaluating commercial building lighting and HVAC systems. High-quality, long-term lighting and end-use metered data were obtained for six commercial buildings.

**Evaluation methods:** Long-term end-use metered data were assembled for each building in the study. A data logger collected true electric power measurements. The data records were averaged over a 15-minute period. A continuous annual time-series data file was assembled for each building. The time-series data records were processed into average daily values for each day of the year. Annual consumption was calculated from the sum of the daily values. Once the actual annual lighting energy consumption was tabulated, the data were segmented into continuous two, three and four week periods. Thus, a series of short-term lighting tests were simulated from the annual time series data. The average daily consumption for weekdays and weekends was calculated for each of the simulated short-term periods, and the annual energy consumption was extrapolated from the daily values for each period. The extrapolated annual consumption was compared to the actual measured annual consumption, thus providing a comparison between the value calculated from a simulated short-term test to the actual value. This exercise was repeated over all possible two, three, and four-week periods throughout the year.

**Findings:** The extrapolation errors associated with short-term monitoring were found to be reasonable. With the exception of one building, the error is generally in the range of 2-8% and the maximum error is in the range of 5-20%. These errors are generally lower than the sampling errors associated with making measurements on a subset of the total lighting fixtures or circuits in a building.

**Source:** Amalfi, J., P. Jacobs, and R. Wright. 1996. "Short-Term Monitoring of Commercial Lighting Systems - Extrapolation from the Measurement Period to Annual Consumption," in the *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*. Vol. 6, pp. 1-7. Washington, D.C.: American Society for an Energy-Efficient Economy.

Short-term monitoring is a useful tool for estimating energy savings when the efficiency of the equipment is enhanced, but the operating hours remains fixed (e.g., constant-load and constant-use equipment, such as hallway lighting and exit signs). Spot metering of the connected load before and after the activity quantifies this change in efficiency with a high degree of accuracy. For activities where the hours of operation are variable, the actual operating (run-time) hours of the activity

should be measured before and after the installation using a run-time meter. Thus, the advantage of the spot meter is that it is simple and easy to apply. This method is more accurate than using engineering calculations, since the parameters are measured instead of being assumed. The primary disadvantage is its limited applicability (i.e., where operating hours are the same before and after treatment). Similar to end-use metering, short-term monitoring is more often seen as a data collection method (rather than a data analysis method) that can provide useful information for integrative methods (see Section 4.2.7).

**Table 7. References to Short-term Monitoring**

Examples	References
Residential weatherization Commercial lighting Industrial process, refrigeration, and miscellaneous measures Paper manufacturing	Bohac et al. (1996) Jacobs et al. (1994) Clarke et al. (1996)  Englander et al. (1996)
	<b>General Reference</b>
	Violette et al. (1991)

#### 4.2.7. Integrative methods (for groups of buildings)

Integrative methods combine one or more of the above methods to create an even stronger analytical tool. These approaches are rapidly becoming the state of the practice in the evaluation field (Raab and Violette 1994). The most common integrative approach is to combine engineering and statistical models where the outputs of engineering models are used as inputs to statistical models (Box 9). These methods are often called Statistically Adjusted Engineering (SAE) methods or Engineering Calibration Approaches (ECA). Although they can provide more accurate results, integrative methods typically increase the complexity and expense. To reduce these costs while maintaining a high level of accuracy, a related set of procedures has been developed to leverage high cost data with less expensive data. These leveraging approaches typically utilize a statistical estimation approach termed ratio estimation that allows data sets on different sample sizes to be leveraged to produce estimates of impacts (see Violette and Hanser 1991). Done properly, ratio estimation will decrease costs because the data needs are less.

**Box 9****Integrative Methods Example**  
(for groups of buildings)

The Pacific Gas and Electric (PG&E) Company contracted with a consulting firm to conduct an integrated and comprehensive evaluation of its Commercial Lighting Program. Two types of data sources were used for the evaluation: Existing data and newly gathered evaluation data. The existing data included PG&E's historical billing data, program participation data, other program-related data, and industry standards information. The new data came from evaluation surveys and metered data. The impact analysis was based on a nested sample design, with a core of lighting-logged sites supplying calibration for the on-site sample, and the on-site audit sample being leveraged with a larger, less expensive, telephone survey. The lighting logger data supplied the most accurate source of data for calibration of engineering estimates. A relatively small on-site auditing sample supported the telephone sample for the largest participation segments. This sample contributed equipment details that were site-specific, and better estimates of operating hours, operating factors, equipment efficiency, lamp burnout rates, etc. The telephone survey supplied information on participant decision-making, energy-related changes at each site for the billing period covered by the billing analysis, etc

Evaluation methods: Demand estimates were based upon engineering models calibrated to on-site data, metered data, and industry standards. The energy impact estimates are derived from a combination of engineering estimates and statistically adjusted engineering (SAE) estimates. In the SAE analysis, engineering estimates are compared to billing data using regression analysis, in order to adjust for behavioral factors of occupants and other unaccounted for effects.

Findings: Gross savings were approximately 300,000 MWh and 63,200 kW. The net savings were approximately 270,000 MWh and 57,000 kW (includes free ridership and participant spillover).

Source: Caulfield, T., and E. Galawish. 1996. "Enlightened Lighting Evaluation: Tightening Up the Process," in the *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*. Vol. 6, pp. 19-26. Washington, D.C.: American Society for an Energy-Efficient Economy.

**Table 8. References to Integrative Methods**  
(for groups of buildings)

Examples	References
Residential new construction	Mahone et al. (1996)
Residential heating, ventilation & air-conditioning	Samiullah et al. (1996)
Commercial heating, ventilation & air-conditioning	Baker et al. (1996)
Commercial lighting	Caulfield and Galawish (1996)
Commercial new construction	Sebold and Wang (1996)
Commercial and industrial	Caulfield and Boertman (1995)
Industrial process, refrigeration, and miscellaneous measures	Clarke et al. (1996)
	<b>General Reference</b>
	Violette et al. (1991)

#### **4.2.8. Application of estimation methods**

Several methods are available for collecting data on energy-efficiency projects: e.g., engineering calculations, surveys, modeling, end-use metering, on-site audits and inspections, and collection of utility bill data. Similarly, several methods are available for evaluating these kinds of projects: e.g., engineering methods, basic statistical models, multivariate statistical models (including multiple regression models and conditional demand models), and integrative methods. If the focus of the monitoring and evaluation is an individual building, then some methods will not be utilized (e.g., basic statistical models, multivariate statistical models, and some integrative methods), since they are more appropriate for a group of buildings.

There is no one approach that is “best” in all circumstances (either for all project types, evaluation issues, or all stages of a particular project). The costs of alternative approaches will vary and the selection of evaluation methods should take into account project characteristics and the kind of load and schedule for the load before the retrofit. As mentioned previously, the load can be constant, variable, or variable but predictable, and the schedule can either be known (timed on/off schedule) or unknown/variable. The monitoring approach can be selected according to the type of load and schedule.

In addition to project characteristics, the appropriate approach depends on the type of information sought, the value of information, the cost of the approach, and the stage and circumstances of project implementation. The applications of these methods are not mutually exclusive; each approach has different advantages and disadvantages (Table 9), and there are few instances where an evaluation method is not amenable to most energy-efficiency measures. Using more than one method can be informative. Employing multiple approaches, perhaps even conducting different analyses in parallel, and integrating the results, will lead to a robust evaluation. Such an approach builds upon the strengths and overcomes the weaknesses of individual approaches. Also, each approach may be best used at different stages of the project life cycle and for different measures or projects. An evaluation plan should specify the use of various analytical methods throughout the life of the project and account for the financial constraints, staffing needs, and availability of data sources.

Finally, in developing countries, some of these methods may be difficult to implement. For example, in Eastern European countries, metering of energy use at the building level is the most common type of energy metering available and not all buildings are metered (Vine and Kazakevicius 1998). And where people live in apartments, metering of individual apartments is almost nonexistent. Utility bill analysis, therefore, would be impractical; field-calibrated engineering analysis would have to be conducted.

**Table 9. Advantages and Disadvantages of Data Collection and Analysis Methods**

<b>Methods</b>	<b>Application</b>	<b>Advantages</b>	<b>Disadvantages</b>
<b>Engineering Methods</b>	Individual buildings and groups of buildings	Relatively quick and inexpensive for simple engineering methods. Most useful as a complement to other methods. Methods are improving. Useful for baseline development.	Relatively expensive for more sophisticated engineering models. Need to be calibrated with onsite data. By themselves, not good for evaluation of spillover.
<b>Basic Statistical Models</b>	Primarily for groups of buildings	Relatively inexpensive and easy to explain.	Assumptions need to be confirmed with survey data and other measured data. Limited applicability. Cannot evaluate peak impacts. Large sample sizes needed.
<b>Multivariate Statistical Models</b>	Primarily for groups of buildings	Can isolate project impacts better than basic statistical models.	Same disadvantages as for basic statistical models. Relatively more complex, expensive, and harder to explain than basic statistical models.
<b>End-use Metering</b>	Individual buildings and groups of buildings	Most accurate method for measuring energy use. Most useful for data collection, not analysis.	Can be very costly. Small samples only. Requires specialized equipment and expertise. Possible sample biases. Difficult to generalize to other projects. Does not, by itself, calculate energy savings. Difficult to obtain pre-installation consumption.
<b>Short-term Monitoring</b>	Individual buildings and groups of buildings	Useful for measures with relatively stable and predictable operating characteristics. Relatively accurate method. Most useful for data collection, not analysis.	Limited applicability. Using this method alone, energy savings cannot be calculated.
<b>Integrative Methods</b>	Primarily for groups of buildings	Relatively accurate.	Relatively more complex, expensive, and harder to explain than some of the other models.

#### 4.2.9. Application of IPMVP approach

In an earlier report, we reviewed several protocols and guidelines that were developed for the MERV of GHG emissions in the energy sectors by governments, nongovernmental organizations, and international agencies (Vine and Sathaye 1997). Although not targeted to carbon emissions, we believe that the U.S. Department of Energy's (DOE) International Performance Measurement and Verification Protocol (IPMVP) is the preferred approach for monitoring and evaluating energy-efficiency projects for individual buildings and for groups of buildings, since the IPMVP covers many of the issues discussed in these guidelines as well as offering several measurement and verification methods for user flexibility (Kats et al. 1996 and 1997; Kromer and Schiller 1996; USDOE 1997).<sup>1</sup> North America's energy service companies have adopted the IPMVP as the industry standard approach to measurement and verification. States ranging from Texas to New York now require the use of the IPMVP for state-level energy efficiency retrofits. The U.S. Federal Government, through the Department of Energy's Federal Energy Management Program (FEMP), uses the IPMVP approach for energy retrofits in Federal buildings. Finally, countries ranging from Brazil to the Ukraine have adopted the IPMVP, and the Protocol is being translated into Bulgarian, Chinese, Czech, Hungarian, Polish, Portuguese, Russian, Spanish, Ukrainian and other languages. When completed, ASHRAE's GPC 14P guidelines will be used to modify the IPMVP (see Section 1.6).

A key element of the IPMVP is the definition of two measurement and verification (M&V) components: (1) verifying proper installation and the measure's potential to generate savings; and (2) measuring (or estimating) actual savings. The first component involves the following: (a) the baseline conditions were accurately defined and (b) the proper equipment/systems were installed, were performing to specification, and had the potential to generate the predicted savings. The general approach to verifying baseline and post-installation conditions involves inspections, spot measurement tests, or commissioning activities.<sup>2</sup>

The IPMVP was built around a common structure of four M&V options (Options A, B, C, and D) (Table 10). These four options were based on the two components to M&V defined above. The purpose of providing several M&V options is to allow the user flexibility in the cost and method of assessing savings. A particular option is chosen based on the expectations for risk and risk sharing

---

<sup>1</sup> The IPMVP is primarily targeted to the monitoring and evaluation of an individual building, in contrast to other protocols (e.g., CPUC 1998) that are aimed at the monitoring and evaluation of programs (involving multiple sites). The protocol can be downloaded via the World Wide Web: <http://www.ipmvp.org>.

<sup>2</sup> Commissioning is the process of documenting and verifying the performance of energy systems so that the systems operate in conformity with the design intent.

between the buyer and seller and onsite and energy-efficiency project specific features. The options differ in their approach to the level and duration of the verification measurements. None of the options are necessarily more expensive or more accurate than the others. Each has advantages and disadvantages based on site specific factors and the needs and expectations of the customer. Project evaluators should use one of these options for reporting on measured energy savings.

DOE is currently revising the IPMVP and is examining how each of the options can be related to the constancy or variations in load and schedule for the load, and the confidence levels of the energy savings associated with each of these options (personal communication from Steve Kromer, Nov. 20, 1998). For example, simple engineering algorithms could be used for projects with constant loads, multivariate statistical models could be used for predictable loads, and more sophisticated engineering models could be used for random (hard to predict) loads. The level of uncertainty in savings will increase as the loads become harder to predict.

**Table 10. Overview of IPMVP's M&V Options**

<b>M&amp;V Options<sup>1</sup></b>	<b>How Savings Are Calculated [reference to LBNL's MERV methods]</b>	<b>Initial Cost<sup>2, 3</sup></b>	<b>Annual Operating Cost<sup>4</sup></b>
<b>Option A:</b> <ul style="list-style-type: none"> <li>Focuses on physical inspection of equipment to determine whether installation and operation are to specification. Performance factors are either stipulated (based on standards or nameplate data) or measured.</li> <li>Key performance factors (e.g., lighting wattage or "motor" efficiency) are measured on a snapshot or short-term basis.</li> <li>Operational factors (e.g., Lighting operating hours or motor runtime) are stipulated based on analysis of historical data or spot/short-term measurements.</li> </ul>	Engineering calculations or computer simulations based on metered data and stipulated operational data.  [Engineering methods (4.2.2)] [Short-term monitoring (4.2.6)]	0.5 to 3%	0.1 to 0.5%
<b>Option B:</b> <ul style="list-style-type: none"> <li>Intended for individual energy conservation measures (ECMs) (retrofit isolation) with a variable load profile.</li> <li>Both performance and operational factors are measured on a short-term continuous basis taken throughout the term of the contract at the equipment or system level.</li> </ul>	Engineering calculations after performing a statistical analysis of metered data.  [Engineering methods (4.2.2)] [End-use metering (4.2.5)]	2 to 8%	0.5 to 3%
<b>Option C:</b> <ul style="list-style-type: none"> <li>Intended for whole-building M&amp;V where energy systems are interactive (e.g. efficient lighting system reduces cooling loads) rendering measurement of individual ECMs inaccurate.</li> <li>Performance factors are determined at the whole-building or facility level with continuous measurements.</li> <li>Operational factors are derived from hourly measurements and/or historical utility meter (electricity or gas) or sub-metered data.</li> </ul>	Engineering calculations based on a statistical analysis of whole-building data using techniques from simple comparison to multivariate (hourly or monthly) regression analysis. [Basic statistical models (4.2.3)] [Multivariate statistical models (4.2.4)]	0.5 to 3% (utility bill analysis)  2 to 8% (hourly data)	0.5 to 3%
<b>Option D:</b> <ul style="list-style-type: none"> <li>Typically employed for verification of saving in new construction and in comprehensive retrofits involving multiple measures at a single facility where pre-retrofit data may not exist.</li> <li>In new construction, performance and operational factors are modeled based on design specification of new, existing and/or code complying components and/or systems.</li> <li>Measurements should be used to confirm simulation inputs and calibrate the models.</li> </ul>	Calibrated energy simulation/ modeling of facility components and/or the whole facility; calibrated with utility bills and/or end-use metering data collected after project completion.  [Engineering methods (4.2.2)] [Integrative methods (4.2.7)]	2 to 8%	0.5 to 3%

**Source:** Adapted from USDOE (1997) and based on personal communication from Greg Katz, USDOE, Dec. 18, 1998.

<sup>1</sup> It is assumed that the cost of minimum M&V, in projects not following IPMVP, involves an initial cost of 0.5%, and an annual operating cost of 0.1% to 0.2%, of the project cost. The costs in this table are uncertain and should be used for general guidance; developers need to estimate costs based on real projects.

<sup>2</sup> The initial M&V cost includes installation and commissioning of meters.

<sup>3</sup> In new construction, this is the % of the difference in cost between baseline equipment and upgraded/more efficient equipment

<sup>4</sup> Annual operating cost includes reporting, data logger and meter maintenance cost over the period of the contract

#### 4.2.10. Quality assurance guidelines

Implementing data collection and analysis methods is both an art and a science, and there is known problems associated with these methods. Thus, simply adhering to minimal standards contained in guidelines is no guarantee that an evaluator is doing a professional job. Accordingly, we have included Quality Assurance Guidelines (QAG) that require evaluators and verifiers to indicate specifically how basic methodological issues and potentially difficult issues were addressed (see Appendices B and C).<sup>1</sup> The guidelines cover key methodological issues associated with each data collection and analysis method.

The QAG should be seen as practice and reporting standards, rather than highly prescriptive methodological standards: the QAG require evaluators to describe how certain key issues were addressed rather than to require them to address these issues in a specific way. Adherence to such guidelines still allows the methods to be shaped by the interaction of the situation, the data, and the evaluator.

The QAG are to be used in three ways. First, they are included in the Monitoring and Evaluation Reporting Form (Appendix B), so that evaluators will know that they will be held accountable for conducting a sound analysis. Second, they are included in the Verification Reporting Form (Appendix B), so that policymakers and other stakeholders could review a verification report and quickly assess whether the evaluator addressed the most basic methodological issues. This is especially important since most stakeholders do not have the time nor the personnel to carefully scrutinize every written evaluation report, let alone attempt to replicate the results of all of these studies. The details of how evaluators addressed these methodological issues should be contained in the very detailed documentation that would be in the technical appendix of any evaluation report, or in working papers. Finally, the QAG can be used to create a common language to facilitate communication among project developers, evaluators, verifiers, policymakers, and other stakeholders.

Evaluators and verifiers should consider the issues involved in conducting these methods, some of which have been described previously, and which are listed in Table 11 and described in more detail in Appendices B and C. The column headings refer to the data collection and analysis methods described in Section 4.2. The rows refer to the types of issues to be considered when addressing each method. Examples of each of these issues are mentioned below:

---

<sup>1</sup> These guidelines are primarily based on the QAG that were developed for the California Demand-Side Management Advisory Committee (CADMAC) (Ridge et al. 1997). In theory, the QAG could be used in the estimation stage, but are not included in the Estimation Reporting Form.

**For individual buildings and groups of buildings:**

- **Calibration:** e.g., were the input assumptions and calculated results of engineering models compared and adjusted to actual data?
- **Data type and sources:** e.g., what was the source of the data and the methods used in collecting data?
- **Outliers:** e.g., how were outliers and influential observations identified and handled?
- **Missing data:** e.g., how were missing data handled?
- **Triangulation:** e.g., if more than one estimate of savings was calculated, how were the results combined to form one estimate?
- **Weather:** e.g., what was the source of weather data used for the analysis?
- **Engineering priors:** e.g., what was the source of prior engineering estimates of savings?
- **Interactions:** e.g., how was the interaction between heating and lighting addressed?
- **Measurement duration:** e.g., what was the duration and interval of metering?

**For groups of buildings:**

- **Sample and sampling:** e.g., what kind of sampling design was used?
- **Collinearity:** e.g., if two or more variables were highly correlated, how were they treated?
- **Specification and error:** e.g., what kind of errors were encountered in measuring variables and how were these errors minimized?
- **Comparison group:** e.g., how was a comparison group defined for estimating net savings?

**Table 11. Quality Assurance Issues for Data Collection and Analysis Methods<sup>1</sup>**

(✓ = applicable; blank = not applicable)

	Engineering Methods	Basic Statistical Models (2)	Multivariate Statistical Models (3)	End-use Metering	Short-term Monitoring	Integrative Methods (4)
Calibration	✓					✓
Data type and sources	✓	✓	✓	✓	✓	✓
Outliers		✓	✓			✓
Missing data		✓	✓	✓	✓	✓
Triangulation			✓			✓
Weather		✓	✓			✓
Engineering priors			✓			✓
Interactions	✓	✓	✓			✓
Measurement duration				✓	✓	✓
Sample and sampling		✓	✓	✓	✓	✓
Specification and error			✓			✓
Collinearity			✓			✓
Comparison group		✓	✓			✓

<sup>1</sup> Quality assurance issues (rows) are described in Appendices B and C, and the data collection and analysis methods are described in Section 4.2

<sup>2</sup> Primarily for analysis of groups of buildings; includes statistical comparison methods

<sup>3</sup> Primarily for analysis of groups of buildings; includes conditional demand analysis models

<sup>4</sup> Primarily for analysis of groups of buildings; includes engineering calibration approaches

#### 4.11. Positive Project Spillover

The methods for estimating positive project spillover are similar to those used for free ridership (Section 4.13.1) (Goldberg and Schlegel 1997; Weisbrod et al. 1994). Explicit estimates can be obtained by asking participants and nonparticipants survey questions, and discrete choice models can be used (e.g., the effect on implementation of program awareness, rather than program participation, is estimated). Participant and nonparticipant spillover effects can be included in savings estimates in billing analyses, similar to how gross savings are calculated (see Box 10).

**Box 10****Project Spillover Example**

A group of utilities in the New England area (New England Electric System, Inc., Boston Edison, Northeast Utilities, Eastern Utilities Associates and Commonwealth Electric System) contracted with a consulting firm to assess the effect of DSM programs on the residential market for compact fluorescent lamps technology and quantify the spillover effects of their residential DSM programs.

Evaluation methods: The study included telephone surveys of participants, nonparticipants and interviews with representatives of major manufacturers of compact fluorescents and retailers, as well as a review of statistical and secondary sources on shipments, sales, and residential saturation of compact fluorescent lamps (CFLs). Three methods were used to estimate spillover: (1) comparison of saturation of CFLs between households in the sponsors' territories and those in nonprogram areas (in the Midwest and South), (2) spillover estimates based on analysis of customer self-reports within the program areas, and (3) discrete choice modeling (which yields estimates of net program savings including spillover and of spillover savings alone).

Evaluation findings: The three methods yielded similar (all within 7% points) net-to-gross ratios. The discrete choice modeling was chosen as the superior methodology, compared to the other two methodologies. The model estimated spillover savings at 27% of gross program savings. The researchers also identified: (1) changes in the behavior of manufacturers which accelerated the market penetration of CFLs; (2) indicators that these changes were likely to persist in the face of the current decline in utility DSM activity; and (3) evidence that the above changes were attributable to utility DSM efforts and, in some cases, to the efforts of the sponsors in particular.

Source: Xenergy, Inc. 1995. *Final Report: Residential Lighting Spillover Study*. Burlington, MA: Xenergy, Inc.

**4.12. Market Transformation**

Most evaluations of market transformation projects focus on market effects (e.g., Eto et al. 1996; Schlegel et al. 1997): the effects of energy-efficiency projects on the structure of the market or the behavior of market actors that lead to increases in the adoption of energy-efficient products, services, or practices. In order to claim that a market has been transformed, project evaluators need to demonstrate the following (Schlegel et al. 1997):

- There has been a change in the market that resulted in increases in the adoption and penetration of energy-efficient technologies or practices.
- That this change was due at least partially to a project (or program or initiative), based both on data and a logical explanation of the program's strategic intervention and influence.
- That this change is lasting, or at least that it will last after the project is scaled back or discontinued.

The first two conditions are needed to demonstrate market effects, while all three are needed to demonstrate market transformation. The third condition is related to the discussion on persistence

(Section 4.1.2): if the changes are not lasting (i.e., they do not persist), then market transformation has not occurred. Because fundamental changes in the structure and functioning of markets may occur only slowly, evaluators should focus their efforts on the first two conditions, rather than waiting to prove that the effects will last.

To implement an evaluation system focused on market effects, one needs to carefully describe the scope of the market, the indicators of success, the intended indices of market effects and reductions in market barriers, and the methods used to evaluate market effects and reductions in market barriers (Schlegel et al. 1997) (see Box 11).

### Box 11

#### Market Transformation Example

The Pacific Gas and Electric (PG&E) Company contracted with a consulting firm to determine the extent to which the current state of the supermarket industry in PG&E's territory reflected the effects of past market interventions by PG&E.

Evaluation methods: Preliminary data collection and analysis activities included a review of PG&E data sources and existing literature; interviews with PG&E program staff; two focus groups within PG&E's service territory and one in the comparison territory served by Commonwealth Edison; a series of open-ended interviews with vendors at the Food Marketing Institute show in Chicago; and an interview with a supermarket specialist. Other primary data collection activities included interviews with PG&E staff, supermarket decision-makers, architects, designers and technical specification managers, and vendors/manufacturers. These primary data collection activities helped to determine how market actions and attitudes were or were not influenced by PG&E's programs. Interviews were designed to elicit both qualitative and quantitative data, and included both open-ended and structured responses.

Evaluation findings: The overall trend in supermarket energy intensity had been downward until 1995, but energy use has been increasing since then. Refrigeration equipment accounts for the largest share (50%) of energy use in this sector. Three manufacturers dominate the refrigeration system industry, while the market for design services is concentrated in a few specialized architects and designers who serve the national market. Local refrigeration contractors supplement in-house supermarket maintenance organizations, playing a critical role in the installation and operation of energy-using equipment. The most fundamental barrier to energy efficiency in the supermarket industry, both now and in the past, is the overwhelming emphasis placed on increasing sales—to the exclusion of energy efficiency and most other operational concerns. In the past several years, barriers to energy efficiency in supermarkets have grown as the result of a number of external forces: marketing, business considerations, regulatory issues, and technology-related concerns. On balance, the PG&E programs appear to have heightened awareness of and interest in energy efficiency; however, supermarkets have become conditioned to expect rebates as a precondition for undertaking energy-efficiency actions. One of the strategies that may help address many of the fundamental barriers to energy efficiency in this industry is to emphasize non-energy benefits in promoting these measures or technologies.

Source: Quantum Consulting, Inc. 1998. *Study of Market Effects on the Supermarket Industry*. Berkeley, CA: Quantum Consulting, Inc.

Evaluation activities will include one or more of the following: (1) measuring the market baseline; (2) tracking attitudes and values; (3) tracking sales; (4) modeling of market processes; and (5) assessing the persistence of market changes (Prahl and Schlegel 1993). As one can see, these evaluation activities will rely on a large and diverse group of data collection and analysis methods, such as: (1) surveys of customers, manufacturers, contractors, vendors, retailers, government organizations, energy providers, etc.; (2) analytical and econometric studies of measure cost data, stocking patterns, sales data, and billing data; and (3) process evaluations.

### **4.3. Re-estimating the Baseline**

During project implementation, the baseline needs to be re-estimated, based on monitoring and evaluation data collected during this period. The re-estimated baseline should describe the existing technology or practices at the facility or site. Ideally, energy use should be measured for at least a full year before the date of the initiation of the retrofit project and for each year after the initiation of the project during the lifetime of the project. However, some types of projects may not require a full year of monitoring prior to the retrofit: e.g., if the loads and operating conditions are constant over time, one-time spot measurement may be sufficient to estimate equipment performance and efficiency.

The monitoring and evaluation of new buildings differs fundamentally from retrofit projects in that existing performance baselines are hypothetical rather than materially existent and are, therefore, generally not physically measurable or verifiable. The implications of this increase with the complexity of measures and strategies to be monitored and verified. The basic steps in the new building monitoring and evaluation do not vary significantly in concept from retrofit monitoring and evaluation.<sup>1</sup>

For new facilities, evaluators often consider the current state or national building code as the baseline.<sup>2</sup> For those states or countries without a building code, standard building practices, usually obtained from builder surveys, are sometimes used as the baseline. However, evaluators should recognize the problems associated with these options (Vine 1996b). The problem with relying on building standards is that builders both exceed and fall below codes. The problem with focusing on builder practices is that the surveys used to characterize building practices may be inaccurate because they are not conducted on a regular basis and rapidly become outdated. Some analysts also

---

<sup>1</sup> The IPMVP has a separate section (Section 6.0) on measurement and verification for new buildings (USDOE 1997).

<sup>2</sup> A few developing countries have building codes (see Janda and Busch 1994).

use the American Society of Heating, Refrigeration and Air-Conditioning Engineers' guidelines, ASHRAE 90.1, as a baseline: ASHRAE 90.1 guides designers in conducting hourly simulations using the specifications in the ASHRAE document. Results obtained from running a simulation on actual buildings is used to determine the level of savings. Simulation results need to be calibrated with actual data to calculate energy savings.

#### **4.3.1. Free riders**

Free ridership can be evaluated either explicitly or implicitly (see Box 12) (Goldberg and Schlegel 1997; Saxonis 1991). The most common method of developing explicit estimates of free ridership is to ask participants what they would have done in the absence of the project (also referred to as "but for the project" discussions). Based on answers to carefully designed survey questions, participants are classified as free riders (yes or no) or assigned a free ridership score. Project free ridership is then estimated as the proportion of participants who are classed as free riders. Two problems arise in using this approach: (1) very inaccurate levels of free ridership may be estimated, due to questionnaire wording;<sup>1</sup> and (2) there is no estimate of the level of inaccuracy, for adjusting confidence levels.

Another method of developing explicit estimates of free ridership is to use discrete choice models to estimate the effect of the program on customers' tendency to implement measures. The discrete choice is the customer's yes/no decision whether to implement a measure. The discrete choice model is estimated to determine the effect of various characteristics, including project participation, on the tendency to implement the measures.

A method for calculating implicit estimates of free ridership is to develop an estimate of savings using billing analysis (as described above) that may capture this effect, but does not isolate it from other impacts. Rather than taking simple differences between participants and a comparison group, however, regression models are used to control for factors that contribute to differences between the two groups (assuming that customers who choose to participate in projects are different from those who do not participate). The savings determined from the regression represent the savings

---

<sup>1</sup> For example, in an analysis of free ridership in a high-efficiency refrigerator program, estimates of free ridership varied from 37% to 89%, depending on questionnaire wording (Boutwell et al. 1992).

associated with participation, over and above the change that would be expected for these customers due to other factors, including free ridership.<sup>1</sup>

### Box 12

#### Free Riders Example

The New England Power Service Company contracted with a consulting firm to estimate free ridership in their commercial/industrial new construction program.

Evaluation method: The study used an extensive survey to probe into what actions participants were likely to have installed in the absence of the program. The survey was administered on-site in 31 facilities to architects, engineers, and designers who had specified 51 different measure installations. An additional 94 such professionals were surveyed by telephone, bringing the total to 125 respondents responsible for 223 measure installations. There was no sampling; attempts were made to reach everyone in the population. Fifty-one nonparticipants were also surveyed. The study used nine general categories: lighting, lighting controls, HVAC equipment, HVAC controls, motors, variable speed drives, refrigeration, building shell, and custom measures. Many survey questions were developed, and the responses to the questions were weighted by the estimated savings accounted for by each project. The purpose of the weighting was to give the larger projects more significance in the calculation of overall free-ridership rates for each measure category.

Evaluation concerns: It was not possible to locate the appropriate respondents for a significant number of installations: e.g., the individual who had worked on a particular project had left the firm or was otherwise not available.

Evaluation findings: Free ridership estimates were: 28-45% for lighting, 62% for lighting controls, 3-9% for HVAC equipment, 16-22% for HVAC controls, 19-80% for motors, 10% for variable speed drives, 0-2% for refrigeration, 90-100% for building shell, and 2-24% for custom measures. Some measure categories (refrigeration, variable speed drives, customer measures, and building shell) had very few respondents and, therefore, provided less confidence in the free-ridership estimates than for other measures. The midpoints of the ranges were used to modify program savings for assessing cost-effectiveness. Significant changes to the program were made as a result of these findings, including the disqualification of building shell measures for financial incentives.

Source: Tokin, B. and G. Reed. 1993. "Free-Ridership Estimation in the New Construction DSM Market," in the *Proceedings of the 1993 Energy Program Evaluation Conference*, pp. 787-791. Chicago, IL: National Energy Program Evaluation Conference.

---

<sup>1</sup> This approach assumes: (1) nonparticipants would naturally buy the energy-efficiency measure as much as participants would, (2) savings from the measures have a significant impact on the bills of nonparticipants, and (3) a sizeable proportion of nonparticipants buy/install the measure. These assumptions are not always valid (personal communication from Adrienne Kandel, California Energy Commission, Jan. 4, 1999).

The U.S. Environmental Protection Agency's Conservation Verification Protocols (Section 1.6.3) reward more rigorous methods of verifying free riders by allowing a higher share of the savings to qualify for tradable SO<sub>2</sub> allowances. Three options are available for verifying free riders: (1) default "net-to-gross" factors for converting calculated "gross energy savings" to "net energy savings;"<sup>1</sup> (2) project-estimated net-to-gross factors, based on measurement and evaluation activities (e.g., market research, surveys, and inspections of nonparticipants) (see Box 13); or (3) if a developer does not do any monitoring nor provide documentation and the default net-to-gross factors are not used, then the net energy savings of a measure will be 50% of the first-year savings, based on one of the methods described in Section 4.2 (Meier and Solomon 1995; U.S. USEPA 1995 and 1996).

### **4.3.2. Comparison groups**

For many projects, comparison groups can be used for evaluating the impacts of energy efficiency projects. Acting as a baseline, comparison groups can capture time trends in consumption that are unrelated to project participation. For example, if the comparison groups' utility bills show an average reduction in energy use of 5% between the pre- and post-periods, and the participants' bills show a reduction of 15%, then it may be reasonable to assume that the estimated project impacts will be 15% minus the 5% general trend for an estimated 10% reduction in use being attributed to the project.

## **4.4. Calculating Net GHG Emissions**

Once the net energy savings have been calculated (i.e., measured energy use minus re-estimated baseline energy use), net GHG emissions reductions can be calculated in one of two ways: (1) if emissions reductions are based on fuel-use or electricity-use data, then default emissions factors can be used, based on utility or nonutility estimates (e.g., see Appendix B in USDOE 1994b)<sup>2</sup>; or (2) emissions factors can be based on generation data specific to the situation of the project (e.g., linking a particular project on an hourly or daily basis to the marginal unit it is affecting). In both methods,

---

<sup>1</sup> The "net-to-gross" factor is defined as net savings divided by gross savings. The gross savings are the savings directly attributed to the project and include the savings from all measures and from all participants; net savings are gross savings that are "adjusted" for free riders and positive project spillover. Multiplying the gross savings by the net-to-gross factor yields net savings.

<sup>2</sup> The emission factors represent the basic conversion between energy consumption and generation of greenhouse gases. These factors are usually expressed in mass of emitted gas per unit of energy input (g/GJ) or sometimes in mass of gas per mass of fuel (g/kg or g/t).

emissions factors translate consumption of energy into GHG emission levels (e.g., tons of a particular GHG per kWh saved). In contrast to default emission factors (method #1), the advantage of using the calculated factors (method #2) is that they can be specifically tailored to match the energy efficiency characteristics of the activities being implemented by time of day or season of the year. For example, if an energy-efficiency project affects energy demand at night, then baseload plants and emissions will probably be affected. Since different fuels are typically used for baseload and peak capacity plants, then emission reductions will also differ.

### Box 13

#### Net-to-Gross Energy Savings Example

The Pacific Gas and Electric (PG&E) Company contracted with a consulting firm to conduct an impact evaluation of its 1994 Industrial Program, specifically industrial process measures (e.g., modifications to food processing systems, oil pumping systems, process boilers, compressors, pumps, dryers, and pollution control equipment). The evaluation was PG&E's largest evaluation to date to employ a "project-specific" engineering approach.

**Evaluation methods:** To determine gross impacts, projects were categorized into evaluation strata based on measure type, measure impact, and project-specific evaluation cost. Large impact projects typically received extensive project-specific engineering approach to determine gross impacts, and smaller impact projects received simple verifications of installation. In general, the evaluation approach consisted of the following steps: (1) verify installation; (2) review and improve on PG&E's impact methodologies; (3) collect post-retrofit data (e.g., actual operating conditions and equipment usage patterns); and (4) measure/monitor key operating parameters.

The net-to-gross analysis was project specific as well, with each project in the evaluation sample receiving a project-specific net-to-gross analysis based on a series of customer interviews (onsite and telephone). For this evaluation, spillover effects were assumed to be small relative to the primary program impacts, and the net-to-gross analysis focused on measuring the impacts of free ridership (four net-to-gross classifications were created).

**Evaluation concerns:** Self-reported data are prone to subjectivity and ambiguity: in practice, the distinction between a free rider and a program-induced participant can frequently be obscure. In many cases, there are elements of both program-induced participation and free ridership in a customer's decision to implement a single energy-efficiency project.

**Evaluation findings:** The net-to-gross analyses showed a high level of free ridership (about 50%). Larger projects had a greater tendency toward free ridership because customers were inclined to identify and implement these projects (for monetary savings and other strategic reasons) independent of motivation from PG&E.

**Source:** Clarke, L., F. Coito, and F. Powell. 1996. "Impact Evaluation of Pacific Gas & Electric's Industrial Process, Refrigeration, and Miscellaneous Measures Programs," in the *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*. Vol. 6, pp. 27-34. Washington, D.C.: American Society for an Energy-Efficient Economy.

The calculations become more complex (but more realistic) if one decides to use the emission rate of the marginal generating plant (multiplied by the energy saved) for each hour of the year, rather than the average emission rate for the entire system (i.e., total emissions divided by total sales) (Swisher 1997). For the more detailed analysis, one must analyze the utility's existing expansion plan to determine the generating resources that would be replaced by saved electricity, and the emissions from these electricity-supply resources. Thus, one would establish a baseline (current power expansion plan, power dispatch, peak load/base load, etc.), select a monitoring domain, conduct monitoring option, measure direct emission reductions (e.g., reductions occurring at the neighboring power plant to lower demand), measure indirect emissions (e.g., modification in the power system due to lower output at the neighboring plant), and calculate net carbon reductions.

One would have to determine if the planned energy-efficiency measures would reduce peak demand sufficiently and with enough reliability to defer or obviate planned capacity expansion. If so, the deferred or replaced source would be the marginal expansion resource to be used as a baseline. This type of analysis may result in more accurate estimates of GHG reductions, but this method will be more costly and require expertise in utility system modeling. In addition, this type of analysis is becoming more difficult in those regions where the utility industry is being restructured: e.g., the supply of energy may come from multiple energy suppliers, either within or outside the utility service area.

The decision on which methodology to use will depend on project size (e.g., kWh, kW, carbon credits requested, project expenditures) or relative project size (e.g., MW/utility service MW). It is up to the evaluator to decide on the best method for the project. Certain thresholds may need to be developed. If a project is of a certain relative magnitude (e.g., a project is 50 MW and the utility's service area is 400 MW), the evaluator should probably select the second method above.